

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日  
Date of Application:

2002年 9月30日

Masato SOME, et al. Q77645  
METHOD OF NORMALIZING GENE.....  
Darryl Mexic 202-293-7060  
September 29, 2003

出 願 番 号  
Application Number:

特願2002-285201

[ ST.10/C ]:

[ JP2002-285201 ]

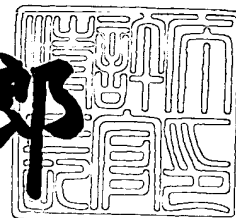
出 願 人  
Applicant(s):

富士写真フイルム株式会社

2003年 4月18日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

太田信一郎



出証番号 出証特2003-3028676

【書類名】 特許願  
【整理番号】 P27126J  
【あて先】 特許庁長官 太田 信一郎 殿  
【国際特許分類】 G01N 33/53  
G06F 19/00

【発明者】

【住所又は居所】 神奈川県足柄上郡開成町宮台 7 9 8 番地 富士写真フイルム株式会社内

【氏名】 染 真人

【発明者】

【住所又は居所】 神奈川県足柄上郡開成町宮台 7 9 8 番地 富士写真フイルム株式会社内

【氏名】 小倉 信彦

【特許出願人】

【識別番号】 000005201

【氏名又は名称】 富士写真フイルム株式会社

【代理人】

【識別番号】 100073184

【弁理士】

【氏名又は名称】 柳田 征史

【選任した代理人】

【識別番号】 100090468

【弁理士】

【氏名又は名称】 佐久間 剛

【手数料の表示】

【予納台帳番号】 008969

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1  
【物件名】 要約書 1  
【包括委任状番号】 9814441  
【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 遺伝子発現データの正規化方法

【特許請求の範囲】

【請求項 1】 第一および第二のサンプルにおける複数の遺伝子に対する発現量を比較する方法において、前記第二のサンプルにおける発現量のデータを正規化する方法であって、

前記第一のサンプルにおける発現量の対数を横軸に、前記第二のサンプルにおける発現量の対数を縦軸にとった対数座標上に、前記第一および第二のサンプルから得られた発現量のデータをプロットにより表示し、

該プロットを傾き 1 の直線により近似して得られる近似直線の前記縦軸における切片の値から係数を求め、

該係数で前記第二のサンプルにおける前記複数の遺伝子に対する発現量のデータを割ることにより該第二のサンプルにおける発現量のデータを正規化する、各工程を含むことを特徴とする遺伝子発現データの正規化方法。

【請求項 2】 第一および第二のサンプルにおける複数の遺伝子に対する発現量を比較する方法において、前記第二のサンプルにおける発現量のデータを正規化する方法であって、

前記第一のサンプルにおける発現量を横軸に、前記第二のサンプルにおける発現量を縦軸にとった座標上に、前記第一および第二のサンプルから得られた発現量のデータをプロットにより表示し、

該プロットを前記座標の原点を通る直線により近似して得られる近似直線の傾きを求め、

該傾きで前記第二のサンプルにおける前記複数の遺伝子に対する発現量のデータを割ることにより該第二のサンプルにおける発現量のデータを正規化する、各工程を含むことを特徴とする遺伝子発現データの正規化方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、複数の遺伝子とハイブリダイズさせることによって得られた 2 つの

サンプルについての遺伝子発現データの比較において、一方のサンプルについての遺伝子発現データを正規化する方法に関する。

【0002】

【従来の技術】

生体内の遺伝情報はDNA塩基配列として保存されており、遺伝子の発現を解析することは各種疾病の予防、早期診断治療、オーダーメイド医療などに有効である。生物学、医学分野での遺伝子解析においては、遺伝子の発現を解析する方法として、マイクロアレイ法が用いられている。マイクロアレイ法は、スライドガラス上に数千個のDNAスポットを作成したDNAチップやDNAマイクロアレイにサンプルをハイブリダイゼーションさせ、ハイブリッド形成の強度を指標にして各遺伝子の発現量を測定する方法である。

【0003】

近年、マイクロアレイ法を使用して遺伝子発現をモニタリングする技術が開発されている。多くの疾病状態は、特定の遺伝子の遺伝的DNAのコピー数の変化または転写レベルの変化によって、様々な遺伝子の発現レベルの違いにより特徴付けられる。例えば、遺伝物質の欠失および獲得は、癌化および癌の進行において重要な役割を担う。また、特定の遺伝子の発現レベルの変化は、様々な癌の存在および進行の指標として機能する。したがって、遺伝子発現の解析においては、疾病を有する細胞と正常な細胞との間で、複数の遺伝子の発現レベルを比較しなければならない。

【0004】

マイクロアレイ法を使用して遺伝子発現の定量を行う際には、検体からの遺伝子抽出量が実験ごとに変化すること等から、基準となる測定値を用いて遺伝子発現データを正規化することが行われている。従来、2つのサンプルにおける複数の遺伝子に対する発現量を比較する方法では、2つのサンプルから必ず発現するある遺伝子を基準プローブとしてマイクロアレイ上に置き、この遺伝子についての発現量が同じであると仮定して、2つのサンプルにおいてこの発現量が一致するようにいずれかのサンプルについての遺伝子発現データを正規化している。また、2つのサンプルにおいてすべての遺伝子についての発現量の総和が等しいと

仮定して、遺伝子発現量の総和が一致するようにいずれかのサンプルについての遺伝子発現データを正規化しているものもある。

【0005】

【発明が解決しようとする課題】

しかしながら、基準プローブを用いた正規化では、サンプルから必ず発現する遺伝子が全体の遺伝子の発現量を代表しているとは限らず、基準プローブについての発現量を一致させても他の遺伝子についての発現量が一定にならないという原理的な問題があり、精度が悪かった。また、遺伝子発現量の総和が等しいと仮定して正規化する方法では、高い発現量を持つ遺伝子についてのデータが支配的となりこのデータの影響を大きく受けるという問題があり、さらにこの方法ではノイズレベルの遺伝子発現量も加算してしまうため、遺伝子発現量は微量であり多くの遺伝子はノイズレベルの発現量であることから誤差が大きくなるという問題もあった。

【0006】

本発明は上記のような従来技術の問題点に鑑みてなされたものであり、2つのサンプルにおける複数の遺伝子に対する発現量を比較する方法において、一方のサンプルについてのデータを適切に正規化する方法を提供することを目的とするものである。

【0007】

【課題を解決するための手段】

本発明の遺伝子発現データの正規化方法は、第一および第二のサンプルにおける複数の遺伝子に対する発現量を比較する方法において、前記第二のサンプルにおける発現量のデータを正規化する方法であって、前記第一のサンプルにおける発現量の対数を横軸に、前記第二のサンプルにおける発現量の対数を縦軸にとった対数座標上に、前記第一および第二のサンプルから得られた発現量のデータをプロットにより表示し、該プロットを傾き1の直線により近似して得られる近似直線の前記縦軸における切片の値から係数を求め、該係数で前記第二のサンプルにおける前記複数の遺伝子に対する発現量のデータを割ることにより該第二のサンプルにおける発現量のデータを正規化する、各工程を含むことを特徴とするも

のである。

【 0 0 0 8 】

また本発明の遺伝子発現データの正規化方法は、第一および第二のサンプルにおける複数の遺伝子に対する発現量を比較する方法において、前記第二のサンプルにおける発現量のデータを正規化する方法であって、前記第一のサンプルにおける発現量を横軸に、前記第二のサンプルにおける発現量を縦軸にとった座標上に、前記第一および第二のサンプルから得られた発現量のデータをプロットにより表示し、該プロットを前記座標の原点を通る直線により近似して得られる近似直線の傾きを求め、該傾きで前記第二のサンプルにおける前記複数の遺伝子に対する発現量のデータを割ることにより該第二のサンプルにおける発現量のデータを正規化する、各工程を含むようにしてもよい。

【 0 0 0 9 】

サンプルとしては、例えば細胞や組織から抽出したDNA、ゲノム等の核酸を用いることができる。第一のサンプルを正常な細胞からのサンプルとし、第二のサンプルを例えば疾病状態などの異常な細胞からのサンプルとすることが好ましいが、これに限定されず、第一のサンプルを異常な細胞からのサンプルとし第二のサンプルを正常な細胞からのサンプルとしてもよく、また第一および第二のサンプルのいずれも異常な細胞からのサンプルとしてもよい。

【 0 0 1 0 】

【発明の効果】

本発明の遺伝子発現データの正規化方法によれば、第一および第二のサンプルにおける複数の遺伝子に対する発現量を比較する場合において、第一のサンプルにおける発現量の対数を横軸に、第二のサンプルにおける発現量の対数を縦軸にとった対数座標上にそれぞれのサンプルにおける発現量のデータをプロットにより表示し、プロットを傾き 1 の直線により近似して得られる近似直線の縦軸における切片の値から係数を求め、この係数で第二のサンプルにおける発現量のデータを割ることにより正規化することとしたので、必ずしも全体の遺伝子の発現量を代表しているとは言えない遺伝子についての発現データを基準とすることにより解析の精度が悪くなることなく、また、高い発現量を持つ遺伝子についての発

現データの影響を大きく受けることなく、マイクロアレイ法で得た2つのサンプルにおける遺伝子発現データを精密に比較し解析することができ、病気の診断や予防、罹病の可能性等を正確に判断することができる。また、複数の遺伝子についての発現量を同時に精密に比較することができるため、2つのサンプルにおいて発現レベルの異なる遺伝子を発見し、例えば病気に冒されているが感染していない個人からのサンプルと感染している個人からのサンプルとの発現結果を比較することにより、その病気に対する抵抗性を与える遺伝子を同定し得る。また、同一の疾病の連続的な段階または進行レベルの組織サンプル間、疾病の最終結果が異なっていたことが知られている組織サンプル間において発現レベルを比較することにより、例えば癌の場合に、悪性細胞と良性細胞との間での発現レベルを比較することができる。

#### 【0011】

また、本発明の遺伝子発現データの正規化方法において、第一および第二のサンプルにおける複数の遺伝子に対する発現量を比較する場合に、第一のサンプルにおける発現量を横軸に、第二のサンプルにおける発現量を縦軸にとった座標上にそれぞれのサンプルにおける発現量のデータをプロットにより表示し、プロットを座標の原点を通る直線により近似して得られる近似直線の傾きを求め、この傾きで第二のサンプルにおける発現量のデータを割ることにより正規化することとした場合も、上記と同様の効果が得られる。

#### 【0012】

##### 【発明の実施の形態】

以下、図面を参照して本発明による遺伝子発現データの正規化方法の実施形態について説明する。

#### 【0013】

図1は、本発明による遺伝子発現データの正規化方法の一実施の形態において、第一および第二のサンプルにおける複数の遺伝子に対する発現量を示すグラフである。図1に示される横軸1は第一のサンプルにおける遺伝子発現量の対数を表し、縦軸2は第二のサンプルにおける遺伝子発現量の対数を表す。マイクロアレイ法により第一および第二のサンプルの両方で発現が測定された遺伝子につい



て、第一および第二のサンプルにおける遺伝子発現量がそれぞれ $x$ 、 $y$ と測定され、それらの対数である $\log x$ 、 $\log y$ が横軸1および縦軸2からなる対数座標上にプロット3により表示される。

【0014】

図1において、複数のプロット3を傾き1の直線により近似して得られる近似直線4を引き、近似直線4の縦軸2における切片の値を $a$ とすると、近似直線4は以下の数式(1)により表される。

【0015】

$$\log y = \log x + a \dots (1)$$

第一および第二のサンプルにおける遺伝子発現データを適切に比較するためには、図1に示す近似直線4を図2に示すように対数座標上の原点を通る直線に補正することが好ましい。この直線は以下の数式(2)により表される。

【0016】

$$\log y' = \log x \dots (2)$$

数式(2)において $\log y' = \log y - a$ であるから、以下の数式(3)が成り立つ。

【0017】

$$y' = y / 10^a \dots (3)$$

したがって、第二のサンプルにおける複数の遺伝子に対する発現量のデータを $10^a$ という係数で割ることにより、第二のサンプルにおける発現量のデータを適切に正規化することができ、第一および第二のサンプルにおける複数の遺伝子に対する発現量のデータを適切に比較することによって病気の診断や予防、罹病の可能性等を正確に判断することができる。

【0018】

本実施の形態における遺伝子発現データの正規化方法においては、プロット3を傾き1の直線により近似することとしているが、1以外の傾きの直線により近似してもよく、また曲線等で近似することとしてもよい。

【0019】

図3は、本発明による遺伝子発現データの正規化方法の第二の実施の形態において、第一および第二のサンプルにおける複数の遺伝子に対する発現量を示すグ

ラフである。図3に示される横軸6は第一のサンプルにおける遺伝子発現量を表し、縦軸7は第二のサンプルにおける遺伝子発現量を表す。マイクロアレイ法により第一および第二のサンプルの両方で発現が測定された遺伝子について、第一および第二のサンプルにおける遺伝子発現量がそれぞれ $x$ 、 $y$ と測定されると、それらが横軸6および縦軸7からなる座標上にプロット9により表示される。

【0020】

図3において、複数のプロット9を座標の原点を通る直線により近似して得られる近似直線10を引き、近似直線10の傾きを $b$ とすると、近似直線10は以下の数式(4)により表される。

【0021】

$$y = bx \dots (4)$$

第一および第二のサンプルにおける遺伝子発現データを適切に比較するためには、図3に示す近似直線10を図4に示すように座標上の原点を通る傾き1の直線11に補正することが好ましい。この直線は以下の数式(5)により表される。

【0022】

$$y' = x \dots (5)$$

数式(5)において、以下の数式(6)が成り立つ。

【0023】

$$y' = y/b \dots (6)$$

したがって、第二のサンプルにおける複数の遺伝子に対する発現量のデータを $b$ という係数で割ることにより、第二のサンプルにおける発現量のデータを適切に正規化することができ、上記と同様の効果を得ることができる。

【0024】

図1および図3において、プロットの分布にばらつきがある場合には分布の多いところの影響を受けて適切な近似直線が引けない場合があるが、この場合には座標をいくつかのブロックに分け、それぞれのブロック内でプロットの数が同じになるようにランダムにプロットを選択するようにしてもよい。この場合には、選択されたプロットを用いて近似直線を作成することで、発現量の多いデータ、発現量の低いデータをすべて考慮した近似直線を作成することができる。

【Q 0 2 5】

また、上記の実施の形態においては、第二のサンプルにおける遺伝子発現データを正規化することとしているが、第一のサンプルにおける遺伝子発現データを正規化するようにしてもよく、また第一および第二のサンプルにおける遺伝子発現データのいずれをも正規化するようにしてもよい。

【図面の簡単な説明】

【図 1】

本発明による遺伝子発現データの正規化方法の一実施の形態において、第一および第二のサンプルにおける複数の遺伝子に対する発現量の対数を示すグラフ

【図 2】

図 1 におけるグラフを補正したグラフ

【図 3】

本発明による遺伝子発現データの正規化方法の第二の実施の形態において、第一および第二のサンプルにおける複数の遺伝子に対する発現量を示すグラフ

【図 4】

図 3 におけるグラフを補正したグラフ

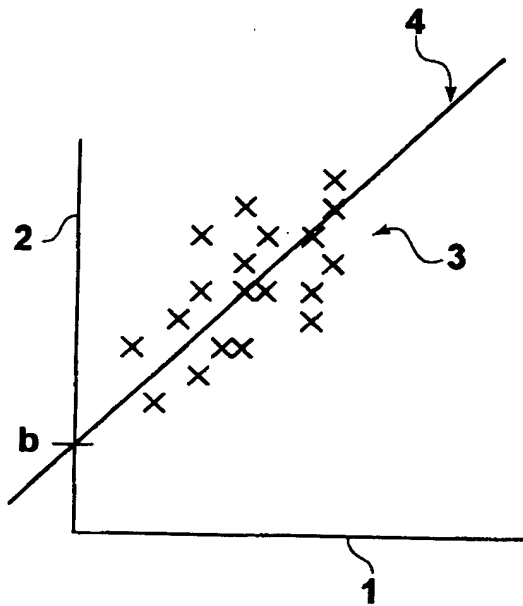
【符号の説明】

- 1、6 横軸
- 2、7 縦軸
- 3、9 プロット
- 4、10 近似直線

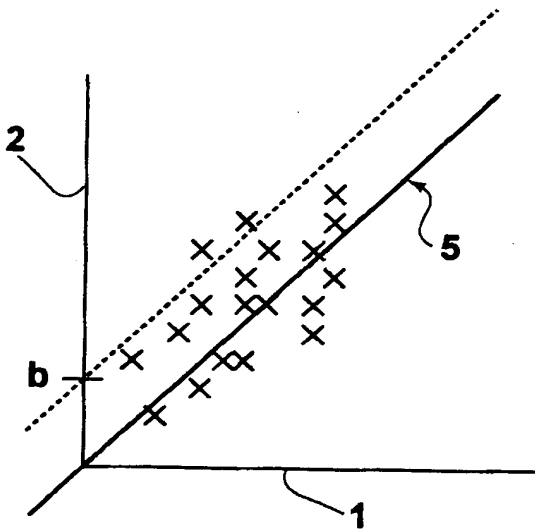
【書類名】

図面

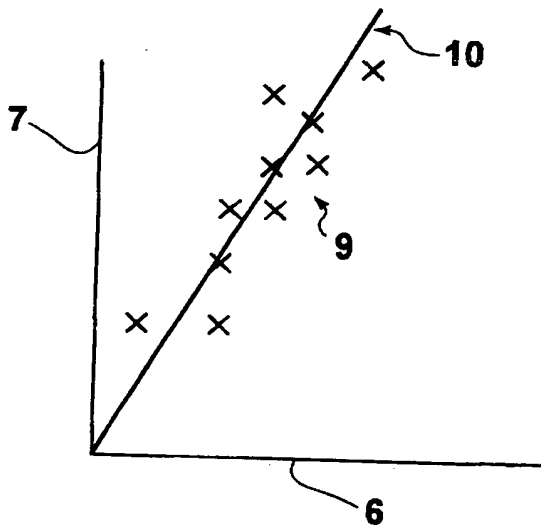
【図1】



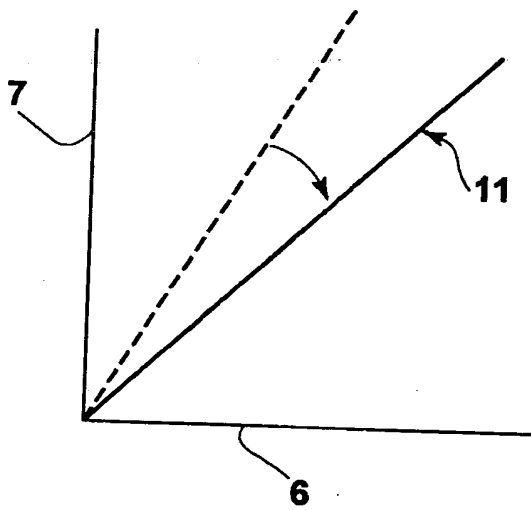
【図2】



【図3】



【図4】



【書類名】 要約書

【要約】

【課題】 第一および第二のサンプルにおける複数の遺伝子発現データの比較において、第二のサンプルにおける遺伝子発現データを適切に正規化する。

【解決手段】 第一のサンプルにおける遺伝子発現量の対数を横軸1に、第二のサンプルにおける遺伝子発現量の対数を縦軸2にとった対数座標上に、第一および第二のサンプルから得られた発現量のデータをプロット3により表示し、プロットを傾き1の直線により近似して得られる近似直線4の縦軸2における切片から求められる係数で第二のサンプルにおける発現量のデータを正規化する。

【選択図】 図 1

認定・付加情報

特許出願の番号	特願 2 0 0 2 - 2 8 5 2 0 1
受付番号	5 0 2 0 1 4 6 3 1 3 6
書類名	特許願
担当官	第一担当上席 0 0 9 0
作成日	平成 1 4 年 1 0 月 8 日

<認定情報・付加情報>

【提出日】	平成14年 9月30日
【特許出願人】	
【識別番号】	000005201
【住所又は居所】	神奈川県南足柄市中沼 2 1 0 番地
【氏名又は名称】	富士写真フイルム株式会社
【代理人】	申請人
【識別番号】	100073184
【住所又は居所】	神奈川県横浜市港北区新横浜 3 - 1 8 - 3 新横 浜 K S ビル 7 階
【氏名又は名称】	柳田 征史
【選任した代理人】	
【識別番号】	100090468
【住所又は居所】	神奈川県横浜市港北区新横浜 3 - 1 8 - 3 新横 浜 K S ビル 7 階
【氏名又は名称】	佐久間 剛

出 願 人 履 歴 情 報

識別番号

[000005201]

1. 変更年月日 1990年 8月14日

[変更理由] 新規登録

住 所 神奈川県南足柄市中沼210番地  
氏 名 富士写真フイルム株式会社